Anderson, Christopher

PREDICT 456 Section 55

Assignment #3

Introduction

While there has been an emphasis throughout professional sports on “bigger, stronger, and faster” performers in recent years, the focus in the NBA has been on players with flexibility. Not flexibility of the body, characterized by a wide range of motion, but rather the ability to play more than position on the court. Based on definitions provided on the NBA’s website, the point guard (PG) is traditionally the smallest and quickest player in the lineup and acts as the “director” of the offense. The shooting guard (SG) is taller than the PG and is responsible for outside shooting and scoring. The small forward (SF) is generally larger than the SG in both height and weight, as he must be able to play inside near the basket, but quick enough to play defense on the perimeter. The power forward (PF) is generally one of the biggest and strongest players on the team. Finally, the center (C) is the tallest player in the lineup. However, in recent years, terms such as “positional flexibility”, “position-less basketball”, and “small ball” have become prominent throughout the sport.

Much has been written about position-less basketball and the evolution of the ideal NBA player. Longtime sportswriter Jeff Zillgitt defines position-less basketball as “having players who can play multiple positions on the court at the same time.” Teams with several interchangeable players are more difficult to defend and can also defend their opponents more aggressively (Zillgitt, 2012). NBA Editor with the Sporting News, Adi Joseph sought insight from several current players and coaches and concluded that teams have been shifting towards putting the five most talented players on the court, regardless of position, since the late 1990s and early 2000s (Joseph, 2016). Chris Herring, a sports journalist for The Wall Street Journal, reported that strictly defined positions have lost relevance in the NBA because young players are learning skills that were traditionally reserved for guards (Herring, 2012). Drew Cannon, a statistician with the NBA’s Boston Celtics, studied conventional position definitions and found a potential market inefficiency that could be gained if teams sought out specific skills rather than specific positions (Cannon, 2010).

Height is no longer the most important factor in measuring an NBA prospect. Investigative reporter David Epstein examined the value of wingspan and discovered that, while most people have a wingspan approximately equal to their height, the average ratio between the two measures in the NBA is 1.06 (Epstein, 2012). He found wingspan to be an important predictor of both blocked shots and rebounds. NBA writer Jonathon Tjarks assets that the ultimate prototype for the “small ball” type of player is Draymond Green of the Golden State Warriors, who is listed at 6 feet 6 inches tall with a wingspan of 7 feet 2 inches. He dribbles, passes, and shoots like a guard, rebounds like a forward or center, and can defend anyone position on the opposing team (Tjarks, 2015). This evolution is not lost on today’s NBA prospects, who aim to avoid being labelled by one particular position, frequently list versatility as one of their strengths, and seek to emulate Green's success (Jones, 2016).

While the research in the prior paragraph makes it clear that NBA prospects are no longer evaluated based on the rigid positional definitions of the past, it offers little quantifiable evidence that can be used to identify these position-less players. The purpose of this exercise is to examine data from the official NBA Draft Combine and create a model to quantify positional flexibility and identify prospects whose physical measurements indicate they may have the ability to play multiple positions.

Methods

Player measurement data from the official NBA Draft Combine were downloaded from the NBA scouting, statistics and analytics service DraftExpress. All publicly available records were included, covering the NBA Draft Combines from 2009 through 2016 and producing 458 observations on 19 variables. Physical measurements such as height, weight, and wingspan were included in the data, as well as the results of various physical tests like the bench press or agility drill. In addition, player statistics data were extracted from Basketball-Reference using the website's Play Index search tool. The search was limited to regular season games from the 2009-10 NBA season through the 2015-16 NBA season, including only the first three seasons within that period for players who debuted in the league between the 2009-10 season and the 2013-14 season. The result was a dataset with 361 observations on 34 variables, one record for each player containing the per-36-minute averages for statistics such as rebounds, assists, and points over their first three NBA seasons combined. The reason for this level of specificity is that NBA teams are looking for a return on an investment made through the NBA draft during the first three seasons of that player's career. Furthermore, as each player's body changes over time, the measurements taken during the NBA Draft Combine become obsolete. The Combine results are likely to be most representative of a player's true athleticism during his first three seasons.

The data from these two sources were merged together, resulting in a dataset of 215 observations on 16 variables. This decrease in observations was caused by the exclusion of players in the Combine dataset who never played in the NBA, as well as players in the NBA Stats dataset who did not attend the official Draft Combine. Many variables were also filtered out in order to focus on key measurements and stats. There were two nominal variables, Name and Position. The values for Position fell into one of five categories – PG, SG, SF, PF, and C – based on the established basketball position model. The remaining fourteen variables contained ratio level data in the form of physical measurements from the Combine and per-36-minute statistics from each player's first three NBA seasons. Height, Weight, Body_Fat, Wingspan, Vertical, Bench, Agility, and Sprint represent measurements taken at the Combine. TRB, AST, STL, BLK, TOV, and PTS represent each player's total rebounds, assists, steals, blocks, turnovers, and points, respectively, per 36 minutes of playing time during his first three seasons. Each variable is defined below in Table 1.

Table 1: Variable Definitions

Variable	Description	Type	Example
Name	Name of player	Nominal	A.J. Price
Height	Height without shoes as measured at NBA Draft Combine (inches)	Ratio	72.50
Weight	Weight of player as measured at NBA Draft Combine (pounds)	Ratio	193
Body_Fat	Body fat percentage as measured at NBA Draft Combine	Ratio	12.4
Wingspan	Distance between fingertips with both arms fully extended to the side as measured at NBA Draft Combine (inches)	Ratio	75.75
Vertical	Maximum vertical jump as measured at NBA Draft Combine (inches)	Ratio	31.00
Bench	Number of bench press reps achieved by player at NBA Draft Combine	Ratio	11
Agility	Time taken to complete the agility drill at NBA Draft Combine (seconds)	Ratio	10.99
Sprint	Time taken to complete the speed drill at NBA Draft Combine (seconds)	Ratio	3.22
TRB	Total rebounds per 36 minutes of playing time during first 3 NBA seasons	Ratio	3.6
AST	Total assists per 36 minutes of playing time during first 3 NBA seasons	Ratio	4.9
STL	Total steals per 36 minutes of playing time during first 3 NBA seasons	Ratio	1.4
BLK	Total blocks per 36 minutes of playing time during first 3 NBA seasons	Ratio	0.1
TOV	Total turnovers per 36 minutes of playing time during first 3 NBA seasons	Ratio	2.3
PTS	Total points per 36 minutes of playing time during first 3 NBA seasons	Ratio	14.6
Position	Position of player according to traditional position labels	Nominal	PG

As the initial step in the exploratory analysis, the descriptive statistics presented in Table 2 were calculated. The mean value of Wingspan was 82.44 and the mean value of Height was 77.72 inches, producing a ratio between the two of 1.06. This figure is in line with the ratio reported by Epstein in 2012. The dataset contained missing values, primarily in Vertical, Bench, Agility, and Sprint, for players who did not participate in those particular drills at the Combine. These missing values were excluded from any subsequent statistical analyses. There were no obvious outliers based on the maximum and minimum values. Out of the 215 observations, there were 43 PGs, 55 SGs, 39 SFs, 41 PFs, and 37 Cs.

Table 2: Descriptive Statistics

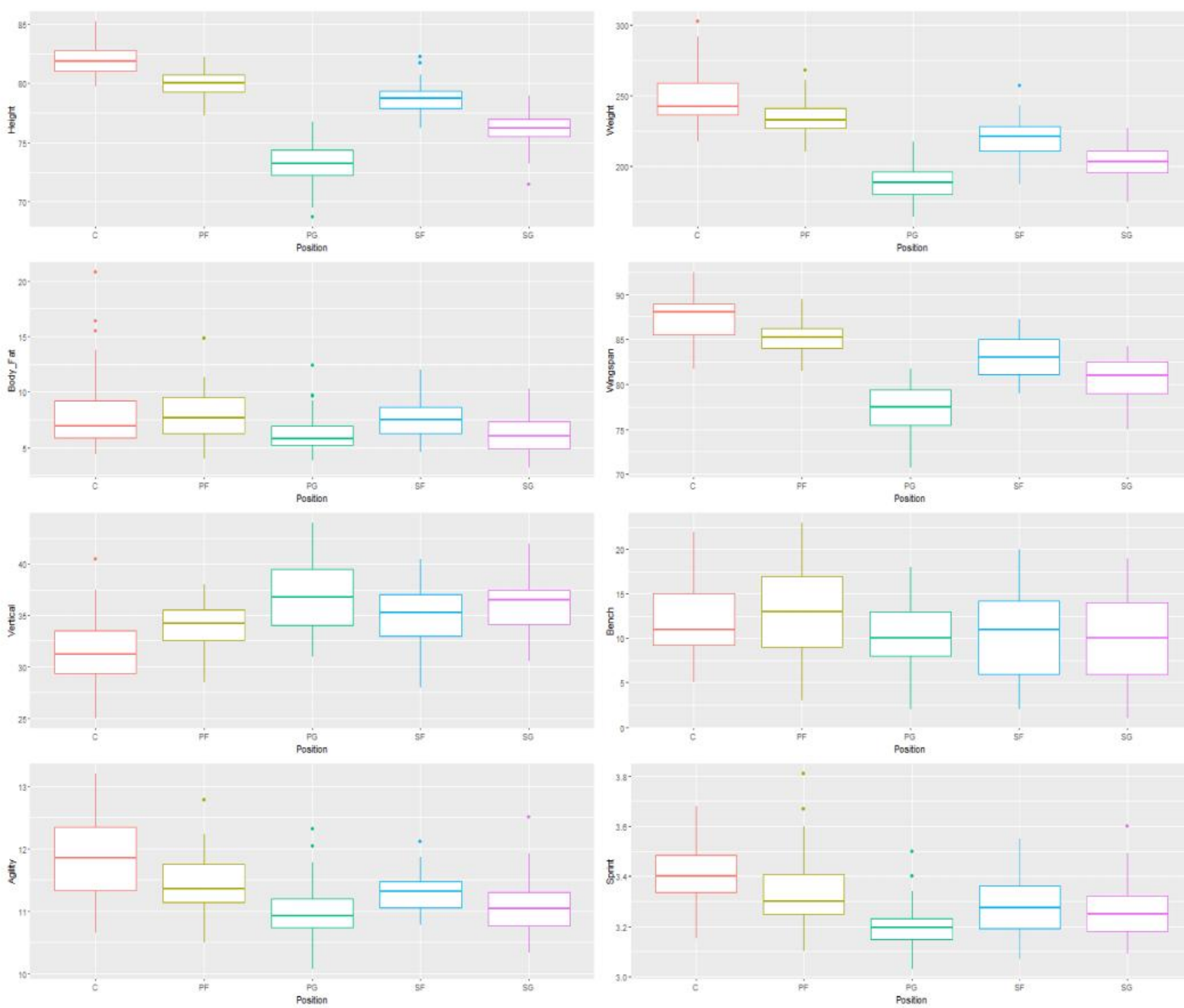
	Height	Weight	Body_Fat	Wingspan	Vertical	Bench	Agility	Sprint
Minimum	68.75	164.00	3.20	70.75	25.00	1.00	10.07	3.030
1st Quantile	75.50	197.20	5.45	79.50	32.50	7.00	10.95	3.190
Median	78.00	216.00	6.70	82.75	35.00	10.00	11.21	3.270
Mean	77.72	216.70	7.15	82.44	34.89	10.89	11.31	3.292
3rd Quantile	80.44	233.00	8.20	85.25	37.00	15.00	11.66	3.380
Maximum	85.25	303.00	20.80	92.50	44.00	23.00	13.20	3.810
NA's	1	1	0	0	17	27	22	18
	TRB	**AST**	**STL**	**BLK**	**TOV**	**PTS**	**Position**	**Count**
Minimum	0.00	0.00	0.00	0.00	0.00	0.00	**C**	37
1st Quantile	3.90	1.40	0.90	0.20	1.60	10.50	**PF**	41
Median	5.70	2.10	1.10	0.50	1.90	12.50	**PG**	43
Mean	6.21	2.72	1.13	0.78	2.10	12.64	**SF**	39
3rd Quantile	8.40	3.65	1.40	1.00	2.40	15.00	**SG**	55
Maximum	15.00	10.30	5.10	4.00	13.70	21.20	**NA's**	0
NA's	0	0	0	0	0	0		

Figure 1: Scatterplot Matrix

Figure 1 on the previous page presents a scatterplot matrix of the continuous variables included in the dataset, along with histogram for each variable and the correlation coefficients for each pair. In general, the histograms for all of the NBA Combine measurements depict a reasonable approximation of the normal distribution. The upper right corner of the graphic, bordered by a dashed line, designates the area where the NBA Combine measurements intersect with the player statistics. In agreement with studies mentioned in the previous section, Wingspan was found to be more strongly correlated than Height to both rebounds (TRB) and blocks (BLK). Interestingly, it was also observed that Vertical produced the only positive correlation to assists (AST) out of the Combine measurements. In addition, none of the physical measurements exhibited a strong correlation to steals (STL), turnovers (TOV), and points (PTS). This suggests that there are factors besides physical attributes, such as timing or shooting ability, involved in producing these statistics.

To further explore the data, a boxplot for each NBA Combine measurement was produced, differentiated by Position. These are displayed below in Figure 2.

Figure 2: Boxplots of NBA Combine measurements by Position



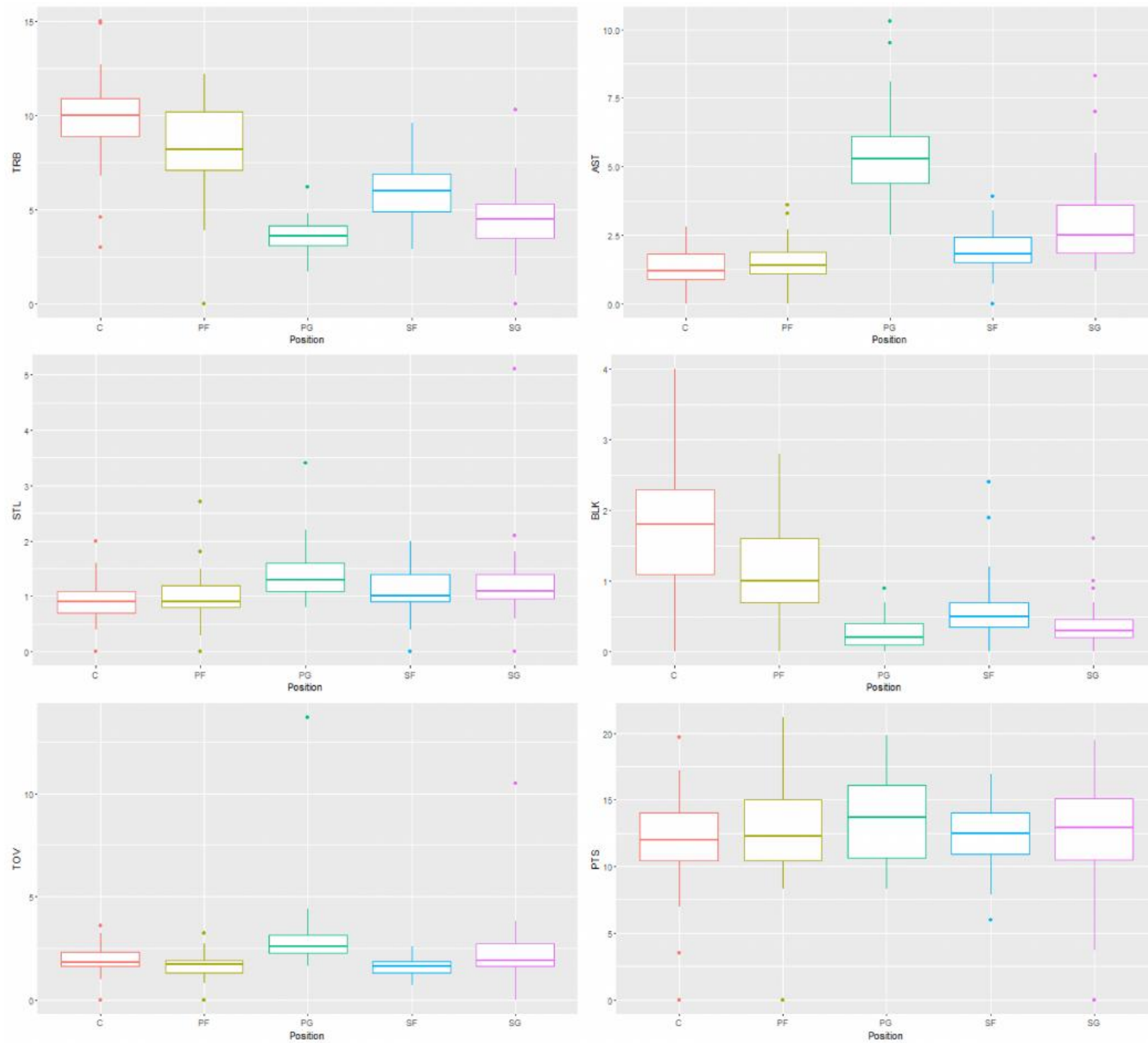
In the boxplots, overlap between the categories suggests that players at multiple positions exhibit similar measurements for the variable in question. There was very little, if any, overlap in the Height plot. This explains why height has traditionally been used to clearly delineate player positions. There was more overlap present in the Wingspan plot, proving, as suggested in the previous section, that it may be an indicator of positional flexibility. There was significant overlap in the plots for Body_Fat, Vertical, Bench, Agility, and Sprint, with the largest differences existing between the PG and C categories on opposite ends of the position spectrum. Therefore, the boxplots indicated that the NBA Combine measurements can be used to identify players with the ability to play multiple positions, rather than classify players into specific positions. Lastly, there were very few outliers present in any of the plots, as defined by the standard boxplot rule which designates any data point that lies far enough away from the interquartile range as an outlier. It was determined that none of the data points required further evaluation for removal from the study.

Next, a boxplot for each player statistic was produced, differentiated by Position. These are displayed in Figure 3 on the following page. These plots were not particularly enlightening, confirming that point guards (PG) and centers (C) have traditionally been relied upon for assists (AST) and blocks (BLK), respectively. However, the overlap in the plots for steals (STL) and points (PTS) show that these important statistics can be accumulated by players at any position and do not require much specialization.

Logistic regression, which estimates the probability that an observation falls into one of two categories, was used to create a model for identifying prospects with the ability to play multiple positions. The first assumption of the binomial logistic regression model is that the dependent variable should be measured on a dichotomous scale. In order to accomplish this, a new binary variable was created for each category in the Position field. The variable names were PG, SG, SF, PF, and C. For example, if the Position value for a particular observation was "PG", then the value under the PG field would be 1. If the Position value was anything other than "PG", it would be 0. A separate logistic regression was run for each binary variable, validating the first assumption. The second assumption is that there are one or more independent variables, which can be either continuous or categorical. Every independent variable used in the model contained continuous data. The third assumption is that each observation is independent. In this study, each observation represented a unique NBA player so the third assumption holds. The fourth assumption is that a linear relationship between the continuous independent variable and the logit transformation of the dependent variable exists. In this case, there is no indication that such a relationship does not exist.

To create the model, a logistic regression was performed on each position-indicating binary variable (PG, SG, SF, PF, and C) with Wingspan, Body_Fat, Vertical, Bench, Agility, and Sprint as the independent variables. Height and Weight were excluded in order to focus on measures that are not traditionally used to define positions. In other words, five separate models were created, one for each position category. Ultimately, the probability of the player being labelled as a PG, SG, SF, PF, or C was estimated for each observation. These five probabilities were summed together to produce a "Flex Score", of which larger values indicate a player who is more likely to possess positional flexibility.

Figure 3: Boxplots of player statistics by Position



Results

Predictive accuracy and predictive value are typically used to assess and interpret a logistic regression model. However, in this study the focus was not prediction or true classification. Rather, the model sought to use the probability from the logistic regression as part of a score used to assess a player's positional flexibility. The top 15 "Flex Scores" are presented below in Table 3.

Table 3: Top 15 Flex Scores

Position probabilities and Flex Scores for all players							
Name	**Position**	**PG_Prob**	**SG_Prob**	**SF_Prob**	**PF_Prob**	**C_Prob**	**Flex_Score**
Hassan Whiteside	C	0.00	0.11	0.10	0.79	0.93	**1.93**
Andre Drummond	C	0.00	0.19	0.26	0.73	0.70	**1.88**
Shane Larkin	PG	1.00	0.76	0.08	0.00	0.00	**1.84**
Jeremy Tyler	C	0.00	0.05	0.64	0.42	0.66	**1.77**
Norris Cole	PG	0.98	0.68	0.06	0.03	0.00	**1.75**
Daniel Orton	C	0.00	0.01	0.49	0.34	0.88	**1.72**
Rudy Gobert	C	0.00	0.07	0.09	0.58	0.98	**1.72**
Phil Pressey	PG	0.98	0.60	0.11	0.01	0.00	**1.70**
Festus Ezeli	C	0.00	0.06	0.21	0.56	0.85	**1.68**
Kemba Walker	PG	0.93	0.56	0.15	0.01	0.00	**1.65**
Isaiah Thomas	PG	0.98	0.51	0.13	0.02	0.00	**1.64**
Andrew Nicholson	PF	0.00	0.09	0.06	0.66	0.82	**1.63**
Ray McCallum	PG	0.90	0.58	0.14	0.01	0.00	**1.63**
Cole Aldrich	C	0.00	0.06	0.28	0.41	0.86	**1.61**
Peyton Siva	SG	0.95	0.52	0.11	0.03	0.00	**1.61**

The table shows several players who may be able to play multiple positions, indicated by the red and orange highlights. Surprisingly, the model indicated that one player, Jeremy Tyler, was nearly an equal fit at both SF and C. However, it was immediately apparent that this list was dominated by point guards and centers, the two positions that reside on opposite ends of the spectrum of NBA player size. Most of the Flex Scores listed above are inflated by a large probability value nearly equal to one in either PG or C. This revelation led to an amendment in the model. As espoused in the Introduction section, position-less basketball is characterized by a lineup filled with players of similar size and athleticism and is regularly referred to as “small ball”. Therefore, players in the first and last quartiles for Height, below 73.5 inches and above 84.4 inches, were removed from the dataset. This left 108 observations that fell into a range of height that is optimal for this type of basketball. The Flex Score model was refit onto the reduced dataset. Table 4 displays a comparison of the two sets of models.

Table 4: Logistic regression model comparison

	All players	**"Small ball" players**
	AIC	**AIC**
PG Model	93.46	14.00
SG Model	189.70	94.15
SF Model	175.80	121.30
PF Model	151.60	69.63
C Model	97.11	14.00

The lower AIC values for the “small ball” models indicate that these are much better models. The top 15 Flex Scores for the “small ball” model are shown below in Table 5.

Table 5: Top 15 Flex Scores for "small ball" players

Position probabilities and Flex Scores for "small ball" players							
Name	**Position**	**PG_Prob**	**SG_Prob**	**SF_Prob**	**PF_Prob**	**C_Prob**	**Flex_Score**
Derrick Caracter	C	0.00	0.00	0.31	0.87	1.00	**2.18**
Michael Carter-Williams	PG	1.00	0.97	0.19	0.00	0.00	**2.16**
Lorenzo Brown	PG	1.00	0.75	0.07	0.01	0.00	**1.83**
Jordan Williams	C	0.00	0.00	0.63	0.12	1.00	**1.75**
Al-Farouq Aminu	SF	0.00	0.03	0.47	0.83	0.00	**1.33**
Derrick Brown	SF	0.00	0.03	0.61	0.69	0.00	**1.33**
Kentavious Caldwell-Pope	SG	0.00	0.87	0.43	0.00	0.00	**1.30**
Quincy Acy	PF	0.00	0.18	0.37	0.75	0.00	**1.30**
Derrick Williams	PF	0.00	0.02	0.64	0.62	0.00	**1.28**
Jeff Adrien	PF	0.00	0.03	0.34	0.91	0.00	**1.28**
Vander Blue	SG	0.00	0.99	0.27	0.00	0.00	**1.26**
Dionte Christmas	SF	0.00	0.75	0.46	0.01	0.00	**1.22**
Patrick Patterson	PF	0.00	0.14	0.21	0.87	0.00	**1.22**
Chase Budinger	SF	0.00	0.63	0.55	0.00	0.00	**1.18**
Alec Burks	SG	0.00	0.80	0.36	0.01	0.00	**1.17**

This table shows several players who may be able to play multiple positions, and also fit into the position-less basketball scheme that has become popular throughout the NBA. There is more flexibility present in the above table, as well, with fewer scores boosted by large values on either end of the position range. Many of these players may not fit the prototypical model for a particular position, but possess physical measurements that substantiate their versatility. The model had the most trouble quantifying the flexibility of small forwards (SF), who are typically the most adaptable player in the lineup already and may not obviously fall into any of the position categories. The complete tables containing all Flex Scores derived from both models are included in Appendix B.

Implications

The logistic regression models produced a score for each player's positional flexibility, or a "Flex Score". This score can be used by NBA teams not as a replacement for scouting, but as a supplement to their existing evaluation process. The model can be tweaked in a variety of ways depending on the type of player a team is seeking. It can focus on smaller players with the ability to play both guard positions (PG and SG) and taller players with the ability to play both "big" positions (PF and C). Not every player must have the ability to play any position in the lineup, like Draymond Green. The flexibility offered by a prospect who can play even two positions successfully is valuable. The model can also focus on players in the "small ball" mold and assess their ability to play a wider range of positions.

Predictive accuracy is not critically essential to the Flex Score model, as it was used merely as a means of identifying versatile prospects and quantifying their positional flexibility. However, the model only assesses players based on physical measurements and athletic performances at the NBA Draft Combine, and could be improved with the inclusion of more variables from other sources of scouting. One obvious basketball skill that is lacking in the model is shooting ability. Shooting is a major factor in the position-less basketball model and, thus, could be even more indicative of positional flexibility than the physical measurements used in this study. Shooting statistics from college or leagues overseas could be included for each prospect.

The model also excludes prospects who did not attend the official NBA Draft Combine. There are many other pre-draft camps and events at which a player can showcase his skills and athleticism. NBA scouting is a comprehensive exercise that must encompass a wide range of factors and should not be limited to data from one event.

Including height and weight in model may have given it more predictive accuracy but that would have undermined the primary purpose of the exercise, which was to build a model that quantifies positional flexibility based on physical measurements besides the traditional height and weight. Adding other variables like shooting could have resulted in a better model for all players rather than focusing on “small ball” players. Despite these limitations, the model deployed in this report provides a framework for quantifying positional flexibility in NBA prospects that could be very useful in pre-draft evaluation.

Appendix A: References

Basketball-Reference NBA & ABA Player Directory. (2000-2016). Retrieved August 14, 2016, from http://www.basketball-reference.com/players/

Cannon, D. (2010, August 2). Five Players. Retrieved August 14, 2016, from http://www.basketballprospectus.com/article.php?articleid=1190

Epstein, D. (2012, November 05). The Case for ... Wingspan. Retrieved August 14, 2016, from http://www.si.com/vault/2012/11/05/106252287/the-case-for--wingspan

Glossary. (2016). Retrieved August 14, 2016, from http://stats.nba.com/help/glossary/

Herring, C. (2012, June 06). The Rise of the Position-Less Player. Retrieved August 14, 2016, from http://www.wsj.com/articles/SB10001424052702303753904577450500740492554

Jones, J. (2016, May 16). NBA prospects don't want to be confined to one position. Retrieved August 14, 2016, from http://www.sacbee.com/sports/nba/sacramento-kings/kings-blog/article77991572.html

Joseph, A. (2016, April 07). LeBron James and Paul George can only avoid NBA's revolution for so long. Retrieved August 14, 2016, from http://www.sportingnews.com/nba/news/lebron-james-paul-george-small-ball-revolution-nba-playoffs-kevin-durant/teauavmfaa4w1cs69qx6n1z8u

NBA Pre-Draft Measurements. (2016). Retrieved August 14, 2016, from http://www.draftexpress.com/nba-pre-draft-measurements/

Player Season Finder. (2000-2016). Retrieved August 14, 2016, from http://www.basketball-reference.com/play-index/psl_finder.cgi

Players and Positions. (2016). Retrieved August 15, 2016, from http://www.nba.com/canada/Basketball_U_Players_and_Posi-Canada_Generic_Article-18037.html

Standing Reach. (2016). Retrieved August 14, 2016, from http://basketball.about.com/od/collegebasketballglossary/g/standingreach.htm

Tjarks, J. (2015, November 20). The Epitome Of Positionless Basketball - RealGM Articles. Retrieved August 14, 2016, from http://basketball.realgm.com/article/239916/The-Epitome-Of-Positionless-Basketball

Zillgitt, J. (2012, October 18). LeBron James, Miami Heat find no position like no positions. Retrieved August 14, 2016, from http://www.usatoday.com/story/sports/nba/heat/2012/10/18/miami-lebron-james-position-versatility-flexibility/1642623/

Appendix B: Expanded Figures

Table 6: Position probabilities and Flex Scores for all players

Position probabilities and Flex Scores for all players							
Name	**Position**	**PG_Prob**	**SG_Prob**	**SF_Prob**	**PF_Prob**	**C_Prob**	**Flex_Score**
Hassan Whiteside	C	0.00	0.11	0.10	0.79	0.93	1.93
Andre Drummond	C	0.00	0.19	0.26	0.73	0.70	1.88
Shane Larkin	PG	1.00	0.76	0.08	0.00	0.00	1.84
Jeremy Tyler	C	0.00	0.05	0.64	0.42	0.66	1.77
Norris Cole	PG	0.98	0.68	0.06	0.03	0.00	1.75
Daniel Orton	C	0.00	0.01	0.49	0.34	0.88	1.72
Rudy Gobert	C	0.00	0.07	0.09	0.58	0.98	1.72
Phil Pressey	PG	0.98	0.60	0.11	0.01	0.00	1.70
Festus Ezeli	C	0.00	0.06	0.21	0.56	0.85	1.68
Kemba Walker	PG	0.93	0.56	0.15	0.01	0.00	1.65
Isaiah Thomas	PG	0.98	0.51	0.13	0.02	0.00	1.64
Andrew Nicholson	PF	0.00	0.09	0.06	0.66	0.82	1.63
Ray McCallum	PG	0.90	0.58	0.14	0.01	0.00	1.63
Cole Aldrich	C	0.00	0.06	0.28	0.41	0.86	1.61
Peyton Siva	SG	0.95	0.52	0.11	0.03	0.00	1.61
Andrew Goudelock	SG	0.76	0.43	0.39	0.01	0.00	1.59
Steven Adams	C	0.00	0.08	0.15	0.62	0.74	1.59
Vander Blue	SG	0.74	0.66	0.17	0.02	0.00	1.59
Darren Collison	PG	0.96	0.42	0.14	0.01	0.00	1.53
Jonny Flynn	PG	0.87	0.51	0.11	0.02	0.00	1.51
Meyers Leonard	C	0.00	0.11	0.09	0.72	0.59	1.51
Solomon Alabi	C	0.00	0.03	0.04	0.46	0.98	1.51
Taj Gibson	PF	0.00	0.09	0.15	0.50	0.76	1.50
Josh Selby	PG	0.66	0.53	0.24	0.01	0.00	1.44
Patrick Mills	PG	0.97	0.31	0.16	0.00	0.00	1.44
Derrick Caracter	C	0.00	0.01	0.11	0.53	0.78	1.43
Ty Lawson	PG	0.99	0.32	0.11	0.01	0.00	1.43
Ekpe Udoh	C	0.00	0.14	0.34	0.43	0.50	1.41
Eric Maynor	PG	0.96	0.36	0.08	0.01	0.00	1.41
Isaiah Canaan	PG	0.85	0.41	0.09	0.05	0.00	1.40
Brandon Knight	PG	0.69	0.49	0.18	0.03	0.00	1.39
Kevin Jones	PF	0.00	0.04	0.11	0.63	0.61	1.39
Greg Monroe	C	0.00	0.02	0.31	0.28	0.77	1.38
Nikola Vucevic	C	0.00	0.05	0.25	0.17	0.91	1.38
Trey Thompkins	C	0.00	0.02	0.44	0.32	0.60	1.38
Larry Sanders	C	0.00	0.07	0.24	0.16	0.90	1.37
Stephen Curry	PG	0.90	0.38	0.07	0.02	0.00	1.37
Michael Carter-Williams	PG	0.50	0.66	0.13	0.05	0.00	1.34
Trey Burke	PG	0.67	0.42	0.23	0.01	0.00	1.33
Jodie Meeks	SG	0.81	0.24	0.22	0.02	0.00	1.29
Henry Sims	C	0.00	0.06	0.19	0.23	0.80	1.28
Avery Bradley	PG	0.54	0.53	0.19	0.01	0.00	1.27
Al-Farouq Aminu	SF	0.00	0.11	0.28	0.42	0.43	1.24
Quincy Acy	PF	0.00	0.30	0.25	0.56	0.13	1.24
Marcus Thornton	SG	0.79	0.32	0.05	0.06	0.00	1.22
Toney Douglas	PG	0.76	0.24	0.18	0.03	0.00	1.21
Derrick Brown	SF	0.00	0.07	0.41	0.43	0.29	1.20
Derrick Favors	PF	0.00	0.12	0.28	0.38	0.42	1.20
Marquis Teague	PG	0.40	0.56	0.20	0.04	0.00	1.20
Nolan Smith	PG	0.73	0.32	0.13	0.02	0.00	1.20
Dewayne Dedmon	C	0.00	0.05	0.26	0.17	0.71	1.19
Jared Sullinger	PF	0.00	0.04	0.05	0.40	0.70	1.19
Jeff Adrien	PF	0.00	0.08	0.20	0.44	0.47	1.19
Derrick Williams	PF	0.00	0.06	0.39	0.47	0.26	1.18
Miles Plumlee	C	0.00	0.21	0.33	0.55	0.04	1.13

Scott Machado	PG	0.85	0.18	0.05	0.05	0.00	1.13
Terrence Ross	SG	0.49	0.54	0.09	0.01	0.00	1.13
Keith Benson	C	0.00	0.15	0.19	0.39	0.38	1.11
Fab Melo	C	0.00	0.06	0.21	0.23	0.60	1.10
Jeff Withey	C	0.00	0.04	0.15	0.19	0.72	1.10
Wayne Ellington	SG	0.57	0.35	0.12	0.05	0.00	1.09
Jarvis Varnado	PF	0.00	0.13	0.36	0.17	0.42	1.08
Tony Mitchell	PF	0.00	0.23	0.08	0.56	0.21	1.08
Darius Johnson-Odom	SG	0.40	0.38	0.16	0.13	0.00	1.07
James Johnson	SF	0.00	0.05	0.45	0.34	0.19	1.03
Rodrigue Beaubois	PG	0.07	0.52	0.40	0.04	0.00	1.03
Jrue Holiday	PG	0.40	0.41	0.17	0.03	0.00	1.01
A.J. Price	PG	0.60	0.09	0.28	0.01	0.00	0.98
Ben McLemore	SG	0.32	0.50	0.14	0.02	0.00	0.98
Cory Joseph	PG	0.43	0.31	0.22	0.02	0.00	0.98
Jeff Teague	PG	0.45	0.35	0.12	0.06	0.00	0.98
Terrence Jones	PF	0.00	0.13	0.17	0.38	0.30	0.98
Tim Hardaway	SG	0.41	0.34	0.10	0.13	0.00	0.98
Jon Leuer	PF	0.01	0.28	0.09	0.53	0.06	0.97
Draymond Green	PF	0.00	0.11	0.31	0.31	0.22	0.95
Thomas Robinson	PF	0.00	0.12	0.26	0.25	0.32	0.95
Victor Oladipo	PG	0.11	0.45	0.17	0.22	0.00	0.95
Brandon Davies	PF	0.00	0.16	0.14	0.33	0.31	0.94
Kentavious Caldwell-Pope	SG	0.28	0.37	0.26	0.03	0.00	0.94
Bernard James	C	0.00	0.20	0.13	0.21	0.39	0.93
JaJuan Johnson	PF	0.00	0.17	0.39	0.28	0.09	0.93
Patrick Patterson	PF	0.01	0.16	0.15	0.40	0.21	0.93
Tony Snell	SG	0.03	0.54	0.17	0.17	0.02	0.93
Quincy Miller	PF	0.00	0.29	0.14	0.39	0.10	0.92
Bradley Beal	SG	0.23	0.48	0.14	0.06	0.00	0.91
Earl Clark	PF	0.00	0.31	0.17	0.22	0.21	0.91
Jordan Crawford	SG	0.37	0.43	0.06	0.05	0.00	0.91
Lazar Hayward	SF	0.00	0.20	0.19	0.43	0.09	0.91
Archie Goodwin	SG	0.11	0.60	0.15	0.04	0.00	0.90
Austin Rivers	SG	0.28	0.47	0.14	0.01	0.00	0.90
Kawhi Leonard	SF	0.00	0.22	0.38	0.07	0.23	0.90
Kevin Murphy	SG	0.34	0.40	0.05	0.11	0.00	0.90
Arnett Moultrie	PF	0.00	0.21	0.37	0.21	0.10	0.89
Enes Kanter	C	0.00	0.14	0.18	0.30	0.26	0.88
Luke Babbitt	SF	0.02	0.32	0.08	0.42	0.04	0.88
Blake Griffin	PF	0.01	0.11	0.16	0.50	0.09	0.87
Damion James	SF	0.00	0.11	0.38	0.23	0.15	0.87
Luke Harangody	PF	0.01	0.02	0.12	0.33	0.39	0.87
Carrick Felix	SG	0.20	0.41	0.08	0.16	0.01	0.86
Jordan Williams	C	0.00	0.02	0.26	0.09	0.49	0.86
Travis Leslie	SG	0.04	0.33	0.40	0.08	0.01	0.86
Tyler Zeller	C	0.01	0.17	0.09	0.45	0.14	0.86
Allen Crabbe	SG	0.03	0.44	0.11	0.24	0.02	0.84
Damian Lillard	PG	0.24	0.39	0.09	0.12	0.00	0.84
Jimmy Butler	SG	0.41	0.23	0.17	0.03	0.00	0.84
Otto Porter	SF	0.00	0.27	0.15	0.31	0.11	0.84
Perry Jones	SF	0.01	0.34	0.24	0.20	0.05	0.84
Scotty Hopson	SG	0.07	0.45	0.18	0.13	0.01	0.84
Tristan Thompson	PF	0.01	0.29	0.23	0.22	0.09	0.84
Jeremy Lamb	SG	0.04	0.53	0.17	0.08	0.01	0.83
Harrison Barnes	SF	0.01	0.19	0.41	0.19	0.02	0.82
Alec Burks	SG	0.09	0.44	0.23	0.04	0.01	0.81
Iman Shumpert	SG	0.11	0.31	0.20	0.19	0.00	0.81
Shelvin Mack	PG	0.31	0.23	0.18	0.09	0.00	0.81
Armon Johnson	PG	0.30	0.25	0.13	0.11	0.01	0.80
Chris Singleton	SF	0.01	0.15	0.41	0.17	0.06	0.80
DaJuan Summers	SF	0.01	0.17	0.26	0.26	0.10	0.80

Doron Lamb	SG	0.27	0.41	0.09	0.03	0.00	0.80
Gani Lawal	PF	0.01	0.07	0.14	0.30	0.28	0.80
Lorenzo Brown	PG	0.39	0.32	0.04	0.04	0.01	0.80
Orlando Johnson	SG	0.02	0.32	0.22	0.22	0.02	0.80
Terrence Williams	SG	0.18	0.38	0.18	0.05	0.01	0.80
Cody Zeller	C	0.07	0.30	0.17	0.23	0.02	0.79
James Southerland	SF	0.00	0.23	0.21	0.20	0.15	0.79
Adonis Thomas	SF	0.00	0.23	0.23	0.27	0.05	0.78
Darius Morris	PG	0.37	0.23	0.12	0.05	0.01	0.78
Charles Jenkins	PG	0.44	0.17	0.10	0.05	0.01	0.77
Chase Budinger	SF	0.20	0.25	0.28	0.04	0.00	0.77
Glen Rice	SG	0.06	0.35	0.26	0.09	0.00	0.76
Tyler Hansbrough	PF	0.01	0.10	0.20	0.33	0.12	0.76
Robert Covington	SF	0.00	0.27	0.14	0.21	0.13	0.75
Dionte Christmas	SF	0.10	0.35	0.25	0.03	0.01	0.74
James Harden	SG	0.02	0.11	0.42	0.16	0.03	0.74
Solomon Hill	SF	0.11	0.27	0.23	0.12	0.01	0.74
Tobias Harris	SF	0.02	0.22	0.35	0.13	0.02	0.74
Wesley Johnson	SF	0.01	0.21	0.22	0.22	0.08	0.74
Mason Plumlee	C	0.03	0.27	0.14	0.26	0.03	0.73
Xavier Henry	SG	0.04	0.38	0.21	0.08	0.02	0.73
Malcolm Thomas	PF	0.01	0.36	0.21	0.10	0.04	0.72
Gerald Henderson	SG	0.10	0.34	0.20	0.05	0.02	0.71
Dante Cunningham	PF	0.04	0.20	0.10	0.30	0.06	0.70
Kenneth Faried	PF	0.01	0.15	0.17	0.27	0.10	0.70
Shabazz Muhammad	SG	0.01	0.24	0.24	0.18	0.03	0.70
Justin Harper	PF	0.02	0.11	0.08	0.28	0.20	0.69
Dominique Jones	SG	0.09	0.16	0.08	0.30	0.05	0.68
Trevor Booker	PF	0.09	0.11	0.24	0.20	0.03	0.67
Erik Murphy	C	0.01	0.06	0.05	0.22	0.32	0.66
Jordan Hill	PF	0.00	0.13	0.19	0.14	0.20	0.66
Ricky Ledo	SG	0.12	0.26	0.24	0.03	0.01	0.66
Mike Muscala	PF	0.00	0.21	0.14	0.14	0.16	0.65
Malcolm Lee	SG	0.12	0.17	0.23	0.10	0.02	0.64
Michael Kidd-Gilchrist	SF	0.01	0.20	0.34	0.04	0.05	0.64
Mike Scott	PF	0.01	0.07	0.16	0.24	0.16	0.64
Willie Warren	PG	0.29	0.21	0.05	0.08	0.01	0.64
Gordon Hayward	SF	0.23	0.18	0.17	0.03	0.01	0.62
Andre Roberson	SG	0.02	0.25	0.16	0.15	0.03	0.61
Devin Ebanks	SG	0.00	0.15	0.15	0.13	0.18	0.61
Evan Turner	SF	0.13	0.22	0.20	0.05	0.01	0.61
Jermaine Taylor	SG	0.13	0.23	0.09	0.15	0.01	0.61
Omri Casspi	SF	0.05	0.25	0.26	0.03	0.02	0.61
Reggie Bullock	SF	0.10	0.26	0.14	0.09	0.01	0.60
Jae Crowder	SF	0.04	0.08	0.12	0.29	0.06	0.59
John Jenkins	SG	0.10	0.28	0.11	0.09	0.01	0.59
Andy Rautins	SG	0.21	0.24	0.04	0.07	0.02	0.58
Chandler Parsons	SF	0.06	0.27	0.18	0.04	0.03	0.58
Craig Brackins	PF	0.01	0.20	0.19	0.11	0.07	0.58
Klay Thompson	SG	0.07	0.22	0.24	0.03	0.02	0.58
Kyle Singler	SF	0.02	0.08	0.33	0.06	0.09	0.58
Tyreke Evans	SG	0.02	0.16	0.32	0.03	0.05	0.58
Danny Green	SG	0.06	0.17	0.10	0.18	0.06	0.57
Darius Miller	SF	0.08	0.21	0.14	0.13	0.01	0.57
Hollis Thompson	SF	0.06	0.32	0.11	0.06	0.01	0.56
James Anderson	SF	0.13	0.11	0.23	0.04	0.02	0.53
Lance Stephenson	SG	0.01	0.13	0.18	0.14	0.07	0.53
Kelly Olynyk	C	0.02	0.21	0.04	0.14	0.10	0.51
Jordan Hamilton	SF	0.03	0.12	0.26	0.05	0.04	0.50
Khris Middleton	SF	0.01	0.20	0.11	0.09	0.09	0.50
Robbie Hummel	SF	0.05	0.07	0.11	0.11	0.08	0.42

Table 7: Position probabilities and Flex Scores for "small ball" players

Position probabilities and Flex Scores for "small ball" players							
Name	**Position**	**PG_Prob**	**SG_Prob**	**SF_Prob**	**PF_Prob**	**C_Prob**	**Flex_Score**
Derrick Caracter	C	0.00	0.00	0.31	0.87	1.00	2.18
Michael Carter-Williams	PG	1.00	0.97	0.19	0.00	0.00	2.16
Lorenzo Brown	PG	1.00	0.75	0.07	0.01	0.00	1.83
Jordan Williams	C	0.00	0.00	0.63	0.12	1.00	1.75
Al-Farouq Aminu	SF	0.00	0.03	0.47	0.83	0.00	1.33
Derrick Brown	SF	0.00	0.03	0.61	0.69	0.00	1.33
Kentavious Caldwell-Pope	SG	0.00	0.87	0.43	0.00	0.00	1.30
Quincy Acy	PF	0.00	0.18	0.37	0.75	0.00	1.30
Derrick Williams	PF	0.00	0.02	0.64	0.62	0.00	1.28
Jeff Adrien	PF	0.00	0.03	0.34	0.91	0.00	1.28
Vander Blue	SG	0.00	0.99	0.27	0.00	0.00	1.26
Dionte Christmas	SF	0.00	0.75	0.46	0.01	0.00	1.22
Patrick Patterson	PF	0.00	0.14	0.21	0.87	0.00	1.22
Chase Budinger	SF	0.00	0.63	0.55	0.00	0.00	1.18
Alec Burks	SG	0.00	0.80	0.36	0.01	0.00	1.17
Kevin Jones	PF	0.00	0.00	0.27	0.90	0.00	1.17
Archie Goodwin	SG	0.00	0.90	0.25	0.01	0.00	1.16
Austin Rivers	SG	0.00	0.86	0.30	0.00	0.00	1.16
Ricky Ledo	SG	0.00	0.64	0.52	0.00	0.00	1.16
Tony Snell	SG	0.00	0.79	0.24	0.12	0.00	1.15
Scotty Hopson	SG	0.00	0.80	0.26	0.08	0.00	1.14
Gani Lawal	PF	0.00	0.06	0.23	0.82	0.00	1.11
Thomas Robinson	PF	0.00	0.05	0.39	0.67	0.00	1.11
Ben McLemore	SG	0.00	0.85	0.25	0.00	0.00	1.10
Glen Rice	SG	0.00	0.62	0.48	0.00	0.00	1.10
DaJuan Summers	SF	0.00	0.20	0.40	0.48	0.00	1.08
Terrence Ross	SG	0.00	0.92	0.16	0.00	0.00	1.08
Terrence Williams	SG	0.00	0.78	0.29	0.01	0.00	1.08
Wayne Ellington	SG	0.00	0.87	0.21	0.00	0.00	1.08
Allen Crabbe	SG	0.00	0.66	0.15	0.26	0.00	1.07
Jeremy Lamb	SG	0.00	0.78	0.26	0.02	0.00	1.06
Gerald Henderson	SG	0.00	0.71	0.31	0.03	0.00	1.05
Damion James	SF	0.00	0.08	0.63	0.33	0.00	1.04
Luke Harangody	PF	0.00	0.01	0.29	0.74	0.00	1.04
Terrence Jones	PF	0.00	0.04	0.31	0.69	0.00	1.04
Solomon Hill	SF	0.00	0.62	0.39	0.02	0.00	1.03
Jarvis Varnado	PF	0.00	0.02	0.65	0.35	0.00	1.02
Kawhi Leonard	SF	0.00	0.16	0.57	0.29	0.00	1.02
Tim Hardaway	SG	0.00	0.84	0.16	0.02	0.00	1.02
James Johnson	SF	0.00	0.02	0.73	0.26	0.00	1.01
Omri Casspi	SF	0.00	0.48	0.52	0.01	0.00	1.01
Dante Cunningham	PF	0.00	0.33	0.14	0.53	0.00	1.00
Justin Harper	PF	0.00	0.11	0.10	0.79	0.00	1.00
Lazar Hayward	SF	0.00	0.16	0.31	0.53	0.00	1.00
Xavier Henry	SG	0.00	0.64	0.32	0.04	0.00	1.00
Draymond Green	PF	0.00	0.04	0.59	0.36	0.00	0.99
Klay Thompson	SG	0.00	0.52	0.46	0.01	0.00	0.99
Tristan Thompson	PF	0.00	0.28	0.35	0.34	0.00	0.97
Evan Turner	SF	0.00	0.54	0.41	0.01	0.00	0.96
Jimmy Butler	SG	0.00	0.67	0.29	0.00	0.00	0.96
Carrick Felix	SG	0.00	0.79	0.10	0.06	0.00	0.95
James Southerland	SF	0.00	0.16	0.39	0.39	0.00	0.94
Tobias Harris	SF	0.00	0.35	0.57	0.02	0.00	0.94
Tony Mitchell	PF	0.00	0.06	0.12	0.76	0.00	0.94
Iman Shumpert	SG	0.00	0.61	0.31	0.01	0.00	0.93
Jared Sullinger	PF	0.00	0.00	0.16	0.77	0.00	0.93
Kevin Murphy	SG	0.00	0.83	0.09	0.01	0.00	0.93
Tyler Hansbrough	PF	0.00	0.10	0.36	0.46	0.00	0.92

Harrison Barnes	SF	0.00	0.24	0.65	0.02	0.00	0.91
Luke Babbitt	SF	0.00	0.39	0.12	0.40	0.00	0.91
Malcolm Lee	SG	0.00	0.46	0.36	0.06	0.00	0.88
Wesley Johnson	SF	0.00	0.24	0.31	0.33	0.00	0.88
Chris Singleton	SF	0.00	0.15	0.61	0.10	0.00	0.86
James Harden	SG	0.00	0.15	0.68	0.03	0.00	0.86
Gordon Hayward	SF	0.00	0.50	0.35	0.00	0.00	0.85
Kenneth Faried	PF	0.00	0.16	0.27	0.42	0.00	0.85
Orlando Johnson	SG	0.00	0.44	0.36	0.04	0.00	0.84
Otto Porter	SF	0.00	0.15	0.26	0.41	0.00	0.82
Kyle Singler	SF	0.00	0.12	0.63	0.06	0.00	0.81
Michael Kidd-Gilchrist	SF	0.00	0.22	0.58	0.01	0.00	0.81
Danny Green	SG	0.00	0.32	0.16	0.32	0.00	0.80
Malcolm Thomas	PF	0.00	0.38	0.36	0.05	0.00	0.79
Reggie Bullock	SF	0.00	0.52	0.26	0.01	0.00	0.79
Tyreke Evans	SG	0.00	0.21	0.56	0.02	0.00	0.79
Trevor Booker	PF	0.00	0.27	0.39	0.12	0.00	0.78
Robert Covington	SF	0.00	0.17	0.21	0.38	0.00	0.76
Shabazz Muhammad	SG	0.00	0.27	0.45	0.04	0.00	0.76
Andy Rautins	SG	0.00	0.60	0.08	0.07	0.00	0.75
Jordan Hamilton	SF	0.00	0.18	0.55	0.01	0.00	0.74
Hollis Thompson	SF	0.00	0.50	0.21	0.01	0.00	0.72
James Anderson	SF	0.00	0.27	0.44	0.01	0.00	0.72
Darius Miller	SF	0.00	0.39	0.27	0.02	0.00	0.68
Jermaine Taylor	SG	0.00	0.47	0.15	0.04	0.00	0.66
Mike Scott	PF	0.00	0.04	0.37	0.24	0.00	0.65
Andre Roberson	SG	0.00	0.29	0.29	0.06	0.00	0.64
Devin Ebanks	SG	0.00	0.08	0.31	0.24	0.00	0.63
Lance Stephenson	SG	0.00	0.12	0.38	0.09	0.00	0.59
Adonis Thomas	SF	0.00	0.12	0.39	0.07	0.00	0.58
Khris Middleton	SF	0.00	0.19	0.25	0.13	0.00	0.57
Jae Crowder	SF	0.00	0.10	0.23	0.21	0.00	0.54
Robbie Hummel	SF	0.00	0.11	0.27	0.13	0.00	0.51

Appendix C: R Code

```
#PREDICT 456 Sports Performance Analysis Section 55 Summer 2016
#Christopher Anderson
#Assignment #3

library(moments)
library(ggplot2)
library(gridExtra)
library(Hmisc)
library(dplyr)
library("psych")

# Read in NBA Combine data downloaded from DraftExpress
combine <- read.csv("NBADraftCombine.csv", header = T, sep = ",")
# 458 observations of 19 variables

# Read in player positions and statistics data downloaded from Basketball-Reference
stats <- read.csv("play-index_psl_finder.cgi_stats.csv", header = T, sep = ",")
# 361 observations of 34 variables

# Merge the datasets together
mydata <- merge(x = combine, y = stats, by.x = "Name", by.y = "Player")
mydata <- subset(mydata,
select=c("Name","Height.w.o.Shoes","Weight","Body.Fat","Wingspan","Max.Vert",

"Bench","Agility","Sprint","TRB","AST","STL","BLK","TOV","PTS","POS"))
colnames(mydata) <-
c("Name","Height","Weight","Body_Fat","Wingspan","Vertical","Bench","Agility","Spri
nt",
                      "TRB","AST","STL","BLK","TOV","PTS","Position")
# 215 observations of 16 variables

# Examine structure of data and summary statistics
str(mydata)
head(mydata)
tail(mydata)
summary(mydata)

# Exploratory scatter plot matrix with correlations and histograms
pairs.panels(mydata[2:15], cex = 1.25, cex.labels = 1.75)

# Exploratory boxplots of combine data by position
plot1 <- ggplot(data = mydata, aes(x = Position, y = Height, group = Position,
                                   colour = Position))+ geom_boxplot(show.legend =
FALSE)
plot2 <- ggplot(data = mydata, aes(x = Position, y = Weight, group = Position,
                                   colour = Position))+ geom_boxplot(show.legend =
FALSE)
```

```
plot3 <- ggplot(data = mydata, aes(x = Position, y = Body_Fat, group = Position,
                                   colour = Position))+ geom_boxplot(show.legend =
FALSE)
plot4 <- ggplot(data = mydata, aes(x = Position, y = Wingspan, group = Position,
                                   colour = Position))+ geom_boxplot(show.legend =
FALSE)
plot5 <- ggplot(data = mydata, aes(x = Position, y = Vertical, group = Position,
                                   colour = Position))+ geom_boxplot(show.legend =
FALSE)
plot6 <- ggplot(data = mydata, aes(x = Position, y = Bench, group = Position,
                                   colour = Position))+ geom_boxplot(show.legend =
FALSE)
plot7 <- ggplot(data = mydata, aes(x = Position, y = Agility, group = Position,
                                   colour = Position))+ geom_boxplot(show.legend =
FALSE)
plot8 <- ggplot(data = mydata, aes(x = Position, y = Sprint, group = Position,
                                   colour = Position))+ geom_boxplot(show.legend =
FALSE)
grid.arrange(plot1, plot2, plot3, plot4, plot5, plot6, plot7, plot8, ncol=2)

# Exploratory boxplots of important statistics by position
plot1 <- ggplot(data = mydata, aes(x = Position, y = TRB, group = Position,
                                   colour = Position))+ geom_boxplot(show.legend =
FALSE)
plot2 <- ggplot(data = mydata, aes(x = Position, y = AST, group = Position,
                                   colour = Position))+ geom_boxplot(show.legend =
FALSE)
plot3 <- ggplot(data = mydata, aes(x = Position, y = STL, group = Position,
                                   colour = Position))+ geom_boxplot(show.legend =
FALSE)
plot4 <- ggplot(data = mydata, aes(x = Position, y = BLK, group = Position,
                                   colour = Position))+ geom_boxplot(show.legend =
FALSE)
plot5 <- ggplot(data = mydata, aes(x = Position, y = TOV, group = Position,
                                   colour = Position))+ geom_boxplot(show.legend =
FALSE)
plot6 <- ggplot(data = mydata, aes(x = Position, y = PTS, group = Position,
                                   colour = Position))+ geom_boxplot(show.legend =
FALSE)
grid.arrange(plot1, plot2, plot3, plot4, plot5, plot6, ncol=2)


# Filter out tallest and shortest quantiles of players to focus on "small ball"
players
# For later use
q25 <- quantile(mydata$Height, probs = 0.25, na.rm = TRUE)
q75 <- quantile(mydata$Height, probs = 0.75, na.rm = TRUE)
q25 <- as.numeric(as.character(q25))
q75 <- as.numeric(as.character(q75))
```

```
flex25 <- factor(mydata$Height < q25, labels = c("yes", "no"))
flex75 <- factor(mydata$Height > q75, labels = c("yes", "no"))
flexdata <- data.frame(mydata, flex25, flex75)
flexdata <- filter(flexdata, flex25 == "yes")
flexdata <- filter(flexdata, flex75 == "yes")

# Create binary variable for each position
PG <- factor(mydata$Position == "PG", labels = c(0,1))
SG <- factor(mydata$Position == "SG", labels = c(0,1))
SF <- factor(mydata$Position == "SF", labels = c(0,1))
PF <- factor(mydata$Position == "PF", labels = c(0,1))
C <- factor(mydata$Position == "C", labels = c(0,1))
mydata <- data.frame(mydata, PG, SG, SF, PF, C)

# Create binary variable for each position for flexdata (for later use)
flexPG <- factor(flexdata$Position == "PG", labels = c(0,1))
flexSG <- factor(flexdata$Position == "SG", labels = c(0,1))
flexSF <- factor(flexdata$Position == "SF", labels = c(0,1))
flexPF <- factor(flexdata$Position == "PF", labels = c(0,1))
flexC <- factor(flexdata$Position == "C", labels = c(0,1))
flexdata <- data.frame(flexdata, flexPG, flexSG, flexSF, flexPF, flexC)

# Logistic regressions for each position using all NBA Combine measurements except
height and weight
PGlog <- glm(PG ~ Body_Fat + Wingspan + Vertical + Bench + Agility + Sprint,
             data=mydata, family = binomial("logit"))
PG_Prob <- predict(PGlog, mydata, type = "response")
SGlog <- glm(SG ~ Body_Fat + Wingspan + Vertical + Bench + Agility + Sprint,
             data=mydata, family = binomial("logit"))
SG_Prob <- predict(SGlog, mydata, type = "response")
SFlog <- glm(SF ~ Body_Fat + Wingspan + Vertical + Bench + Agility + Sprint,
             data=mydata, family = binomial("logit"))
SF_Prob <- predict(SFlog, mydata, type = "response")
PFlog <- glm(PF ~ Body_Fat + Wingspan + Vertical + Bench + Agility + Sprint,
             data=mydata, family = binomial("logit"))
PF_Prob <- predict(PFlog, mydata, type = "response")
Clog <- glm(C ~ Body_Fat + Wingspan + Vertical + Bench + Agility + Sprint,
            data=mydata, family = binomial("logit"))
C_Prob <- predict(Clog, mydata, type = "response")
PGlog  # Residual Deviance: 79.46   AIC: 93.46
SGlog  # Residual Deviance: 175.7   AIC: 189.7
SFlog  # Residual Deviance: 161.8   AIC: 175.8
PFlog  # Residual Deviance: 137.6   AIC: 151.6
Clog   # Residual Deviance: 83.11   AIC: 97.11
results <- data.frame(mydata[c(1,16)], PG_Prob, SG_Prob, SF_Prob, PF_Prob, C_Prob)
results$PG_Prob <- round(as.numeric(as.character(results$PG_Prob)),2)
results$SG_Prob <- round(as.numeric(as.character(results$SG_Prob)),2)
results$SF_Prob <- round(as.numeric(as.character(results$SF_Prob)),2)
```

```
results$PF_Prob <- round(as.numeric(as.character(results$PF_Prob)),2)
results$C_Prob <- round(as.numeric(as.character(results$C_Prob)),2)
Flex_Score <- results$PG_Prob + results$SG_Prob + results$SF_Prob + results$PF_Prob
+ results$C_Prob
results <- data.frame(results, Flex_Score)
View(results)

# Logistic regressions for each position using all NBA Combine measurements except
height and weight using flexdata
PGlog <- glm(flexPG ~ Body_Fat + Wingspan + Vertical + Bench + Agility + Sprint,
             data=flexdata, family = binomial("logit"))
PG_Prob <- predict(PGlog, flexdata, type = "response")
SGlog <- glm(flexSG ~ Body_Fat + Wingspan + Vertical + Bench + Agility + Sprint,
             data=flexdata, family = binomial("logit"))
SG_Prob <- predict(SGlog, flexdata, type = "response")
SFlog <- glm(flexSF ~ Body_Fat + Wingspan + Vertical + Bench + Agility + Sprint,
             data=flexdata, family = binomial("logit"))
SF_Prob <- predict(SFlog, flexdata, type = "response")
PFlog <- glm(flexPF ~ Body_Fat + Wingspan + Vertical + Bench + Agility + Sprint,
             data=flexdata, family = binomial("logit"))
PF_Prob <- predict(PFlog, flexdata, type = "response")
Clog <- glm(flexC ~ Body_Fat + Wingspan + Vertical + Bench + Agility + Sprint,
            data=flexdata, family = binomial("logit"))
C_Prob <- predict(Clog, flexdata, type = "response")
PGlog  # Residual Deviance: 1.263e-07     AIC: 14
SGlog  # Residual Deviance: 80.15   AIC: 94.15
SFlog  # Residual Deviance: 107.3   AIC: 121.3
PFlog  # Residual Deviance: 55.63   AIC: 69.63
Clog   # Residual Deviance: 2.984e-09     AIC: 14
flexresults <- data.frame(flexdata[c(1,16)], PG_Prob, SG_Prob, SF_Prob, PF_Prob,
C_Prob)
flexresults$PG_Prob <- round(as.numeric(as.character(flexresults$PG_Prob)),2)
flexresults$SG_Prob <- round(as.numeric(as.character(flexresults$SG_Prob)),2)
flexresults$SF_Prob <- round(as.numeric(as.character(flexresults$SF_Prob)),2)
flexresults$PF_Prob <- round(as.numeric(as.character(flexresults$PF_Prob)),2)
flexresults$C_Prob <- round(as.numeric(as.character(flexresults$C_Prob)),2)
Flex_Score <- flexresults$PG_Prob + flexresults$SG_Prob + flexresults$SF_Prob +
flexresults$PF_Prob + flexresults$C_Prob
flexresults <- data.frame(flexresults, Flex_Score)
View(flexresults)


# Save datasets for future use
write.csv(mydata, file = "assignment3_mydata.csv")
write.csv(flexdata, file = "assignment3_flexdata.csv")
write.csv(results, file = "assignment3_results.csv")
write.csv(flexresults, file = "assignment3_flexresults.csv")

# End
```